TIME SERIES ANALYSIS OF HOURLY GLOBAL HORIZONTAL SOLAR RADIATION

J. M. Gordon* and T. A. Reddy†

Abstract—Accurate design and optimization of short response time solar energy systems with storage are sensitive to the stationary and sequential characteristics of hourly solar radiation. We perform monthly time series analyses of hourly global horizontal solar radiation for a wide range of climatic stations that span temperate and tropical conditions. The stationary statistics for individual hours are found to be very similar to the corresponding results for daily total global horizontal radiation, in keeping with a related fundamental observation of Liu & Jordan. Investigation of sequential properties shows that autocorrelation coefficients are, to a good approximation, independent of time of day and that persistence times are nearly as long as the entire daylight period, mainly due to the effect of very strong correlations at one-hour lag times. The isolated effect of two-hour and longer lag times, via the partial autocorrelation coefficients, is found to be negligible in most, but by no means all, instances. Finally, we find no universal correlation between hourly autocorrelation coefficients and monthly average radiation figures.

1. INTRODUCTION

Accurate and detailed information on both the stationary and sequential properties of hourly solar radiation are particularly important in the design and optimization of solar energy systems with a "memory" or, equivalently, storage size, of one to a few hours. For systems with very long memory, typical of stand-alone systems, daily total solar radiation data are sufficient since storage effectively smooths out the detailed changes and correlations in hourly radiation. For systems with intermediate memory, say of the order of half a load-day, typical of most solar thermal designs for hot water and space heating applications, storage is sufficiently large that hour-to-hour variations in radiation need not be treated accurately, and day-to-day variations in daily total radiation are on a time scale too long to affect system performance significantly.

Predicting the performance of systems with relatively short memory, however, requires accurate information on the stationary and sequential characteristics of hourly solar radiation. Examples include solar electricity generating systems with buffer storage[1,2], certain solar absorption air conditioning systems[3], and certain control problems in solar buildings[4–6]. (Since in this article we consider the characteristics of hourly global radiation only, we have chosen the references immediately above for systems that utilize global radiation, in contrast to high concentration solar collectors that effectively utilize the beam component only.) The design and analysis of many of these systems, however, were of a one-of-a-kind, custom-designed nature and were performed with large-scale simulations that required extensive climatic databases.

Analytic design techniques have been developed that can obviate these expensive requirements[7,8] but necessitate specific information on the statistics of hourly solar radiation, preferably in as universal a form as possible. The statistical analyses presented in this article are in the spirit of providing the kind of hourly solar radiation input required for such analytic models. They are, however, only a modest first step since they deal with global horizontal radiation only.

There are only a handful of studies that explicitly consider the sequential properties of hourly solar radiation[9–12]. Most studies to date have analyzed solar radiation data on a daily, or longer, time scale[13 and references therein]. This is due mainly to the fact that the database for daily total global horizontal radiation is far larger than that for hourly radiation.

Analyses of the sequential characteristics of assorted meteorological variables, such as cloud cover[14], air temperature, wind speed, and relative humidity[15,16], reveal very strong correlations at the one-hour time lag. They furthermore show that the isolated effect of one-hour lags is the predominant, if not the sole, factor in determining correlations, in contrast to the isolated effect of two-hour (or longer) lag times. We will investigate whether this observation applies to hourly global horizontal solar radiation as well. The sequential properties of solar radiation are expected to be similar to those of the meteorological variables cited above due to the strong causal relationship among them.

In principle, one is interested in the same statistical properties for hourly radiation as have been analyzed for daily radiation. Hence the formalisms developed for the more extensive analyses of daily radiation are applied directly to our studies here.

A recent article[13] analyzed the stationary and
sequential characteristics of daily global horizontal solar radiation for a wide range of climatic conditions. The basic approach and statistical analyses of that paper are applied here to the same database of varying climates, but for the corresponding hourly radiation values (i.e., individual hours of the day).

Our analyses of hourly global horizontal radiation are performed on a *monthly* basis for two reasons. First, long-term average climatic data, as well as energy demand data, are typically tabulated on a monthly basis. Second, one month is a sufficiently short time that the yearly and seasonal deterministic trends can be approximated as constant[22]. The monthly time scale has also been used in analyses of other meteorological variables[14,15]. The major drawback in performing monthly analyses is that it may turn out, for a specific location, that longer time scales can be chosen during which deterministic trends can be well approximated as constant, with the consequent increase in accuracy of a larger statistical sample.

Two important dimensionless variables appear in our analyses that should be defined at the outset. One is the “clearness index,” or ratio of horizontal global (terrestrial) radiation to horizontal extraterrestrial radiation. The symbols $k$ and $K$ will denote clearness index calculated on an hourly and daily basis, respectively. In our analyses of solar radiation for a specific hour, $i$, we employ the variable $x_i = k_i/\bar{k}$, where $\bar{k}$ is the monthly average value of $k_i$ which enables us to compare different hours on a common basis[13].

We also refer to the daily analog of $x_i$, $X = K/\bar{K}$, where $\bar{K}$ is the monthly average value of $K$. In order to retain statistical information on year-to-year fluctuations in clearness index, we calculate the stationary statistics of $x_i$ (and $X$) for each individual year and then average over all years[13], which we refer to as the “grand average” and denote by the brackets $\langle \rangle$.

Our database is summarized in Table 1. It represents a selected sample of the broader database originally analyzed in[13], so as to encompass climatic conditions that span the extremes of clear to cloudy and tropical to temperate. For each entry in Table 1, we note $\langle \bar{k} \rangle$, which is the commonly tabulated characteristic for a climatic station and the grand average of the monthly variance of $X$, $\langle \sigma^2(X) \rangle$.

## 2. HOURLY VS. DAILY STATIONARY STATISTICS

Liu and Jordan[17] observed that the monthly cumulative frequency distribution (CFD) for individual hourly and daily total solar radiation were nearly identical, provided that radiation values were scaled by their mean monthly value. This observation provides an avenue for markedly reducing the time and data requirements in the design of many solar energy systems. Essentially, the CFD for any hour can be obtained from the CFD for daily total radiation only, the latter being far more readily available.

The fundamental distribution of stationary statistics, however, is the Probability Density Function (PDF). Since the CFD is the integrated PDF, it is a less stringent criterion than the PDF for testing hypotheses.

![Image](image.png)

Toward testing the validity of Liu and Jordan’s observation for the PDF, several illustrative cases are shown in Fig. 1 for the new, expanded data base listed in Table 1. In Fig. 1, note that our random variable, $x_i$, is radiation scaled by its monthly average for the particular year, whereas Liu and Jordan used radiation scaled by its multiyear monthly average. The results tend to substantiate Liu and Jordan’s claims, certainly for the key hours of the day that represent the vast majority of daily radiation. Deviations increase as the hour moves farther away from solar noon since (1) day-to-day fluctuations tend to be larger in the early morning and late evening daylight hours as opposed to the hours near solar noon and (2) differences in air mass effects even over a period of one month can be nonnegligible for one or two hours just after sunrise and before sunset.

Because of the strong similarity of the PDF’s for individual hours and daily total radiation, it is not surprising that the observations of[13] for daily total radiation regarding correlations among fundamental
Fig. 1. Databased probability density function for individual hours ($P(x_i)$ vs. $x_i$) and daily total radiation ($P(X)$ vs. $X$).
statistical variables \((\bar{h}_i), (\sigma^2(x_i))\) and grand average skewness \((S(x_i))\) are also found to pertain to individual hourly radiation. Specifically, we confirm that for radiation for individual hours, there is:

1. A trend between \((\bar{h}_i)\) and \((\sigma^2(x_i))\), but no universal correlation, with a statistically significant distinction between temperate and tropical climates; and
2. A correlation between \((\bar{h}_i)\) and \((S(x_i))\) independent of the climate being temperate or tropical, this correlation being nearly identical to that found for daily total radiation\([13]\). The graphical illustration of these points is sufficiently similar to the corresponding graphs of \([13]\) that, in the spirit of economy, they will not be presented here, but can be found in \([18]\).

Finally, the results illustrated in Fig. 1 indicate that the universal analytic functional form proposed in \([13]\) for the PDF of daily total global horizontal radiation can also be used for predicting the PDF's for individual hours. Required input parameters (grand average variance) therefore pertain to daily total radiation data only, which are typically far more readily available than data for individual hours.

### 3. SEQUENTIAL PROPERTIES

Our objective here is to generate and examine the hour-to-hour variations in solar radiation, their persistence times, and persistence strengths. Because different individual hours of the day can have different stationary statistics, it is not admissible to simply determine the sequential characteristics of the variable \(x_i\). Rather, we employ the transformation to the standardized random variable \(z_i\):

\[
z_{i}(j,r,t) = \left(\frac{x_{i}(j,r,t) - \bar{x}(j,r,t)}{\sigma(x_{i}(j,r,t))}\right)
\]

and construct the following weak stationary sequence for each hourly period of the day, \(i\), and each month of the year, \(r\):

\[
z_{i}(1,r,1), z_{i}(2,r,1), \ldots z_{i}(J,r,1), z_{i}(1,r,2),
\]

\[
\ldots z_{i}(J,r,2), \ldots z_{i}(1,r,T), \ldots z_{i}(J,r,T),
\]

The sequential behavior of hourly radiation can be expressed in terms of autocorrelation coefficients\([19,20]\). Except for \([11]\), where the autocorrelation coefficients were generated for each different hour period of the day, all other relevant studies have computed directly the hourly correlations by averaging over all hours of the day\([9,10,12]\). Therefore, we first deemed it appropriate to examine whether the sequential characteristics of hourly radiation are dependent on time of day.

The autocorrelation coefficients, \(\rho(d)\), for lag times of \(d\) hours, were generated for each of the hourly periods of the day and for each of the locations and months listed in Table I, following:

\[
\rho(d) = \left(\frac{1}{JT}\right) \sum_{i=1}^{T} \sum_{j=1}^{J} z_{i}(j,r,t) z_{i-j}(j,r,t).
\]

We note from the results that there is no significant similarity among the various climates. This is illustrated in Fig. 2, which plots the hourly variation of \(\rho(d)\) for three markedly different climatic conditions, and for \(d = 1\) to 9. The fact that there appears to be no universally valid common behavior with respect to climate will be strengthened below. A second important point is that the autocorrelation coefficient for a given time lag, \(d\), can most often be approximated as a constant average value with respect to hour of the day. The variation in \(\rho(d)\) with hour \(i\) is, for most
Global horizontal solar radiation

Fig. 3. Illustrative plots of daily average hourly autocorrelation coefficient at lag time of \( d \) hours, \( \hat{\rho}(d) \), vs. \( d \), including error bars for the standard deviation at the 95\% confidence level.

Fig. 4. Illustrative plots of the partial autocorrelation coefficient, \( \phi_{1}(d,d) \), vs. time of day for lag time of one, two, and three hours. The average over all hours and the associated error bars (for the mean \( \phi_{1}(d,d) \)) at the 95\% confidence level are shown for lag times of one and two hours.
Isolated multihour correlations, in the form of $\phi(d,d)$, are generally small, although they do occasionally exhibit dramatic variations for certain hours of the day.

Like the autocorrelation coefficients $\phi(d)$, $\phi(2,2)$ can be treated, to a good approximation, as being constant and independent of hour of the day, at a significance level of .05. In an identical analysis for all the locations and months listed in Table 1, we found that $\phi(2,2)$ can be assumed to be zero in the large majority of, but by no means all, instances, at a significance level of .05.

This observation regarding two-hour lag times is relevant for two reasons. One is that the strong persistence pattern over nearly the entire daylight period seems to be a consequence of isolated one-hour lag effects only. The other is that there is an important practical implication for the generation of synthetic sequences of hourly solar radiation\[13,19,20\]. Specifically, first-order autoregressive models (or second-order autoregressive models when $\phi(2,2)$ cannot be assumed to be zero), with a Gaussian mapping procedure, may be adequate for the generation of synthetic sequences that accurately capture the sequential characteristics of the actual data set, yet offer the advantage of being summarized in a handful of synoptic parameters and are hence easily transportable.

Finally, is there a generalized correlation between the key autocorrelation coefficient, $\hat{\rho}(1)$, and $\langle K \rangle$—a correlation that would offer a powerful predictive tool in design calculations? The data presented in Fig. 5 illustrate that no such correlation exists, and that $\hat{\rho}(1)$ is large irrespective of $\langle K \rangle$.

In order to ensure that we had not overlooked a more subtle correlation that depended on distinguishing between clear and cloudy periods\[11\], we repeated our analyses separately for the two categories of clear and cloudy days. However, no appreciable improvement in the trend and correlation patterns illustrated above was found.

### 4. CONCLUSIONS

Besides the point of pure meteorological interest, analysis of the stationary and sequential characteristics of hourly solar radiation is important in the optimal design of an important class of solar energy systems, most notably systems with response times of the order of one to a few hours. Accurate solar system designs require not only a knowledge of the frequency of occurrence of various solar radiation levels for each hour of the day but the persistence times and persistence strengths of hour-to-hour variations in solar radiation as well. After selecting an appropriate variable for analysis, namely $x_i = \frac{k_i}{\langle k_i \rangle}$, and the suitable monthly time scale which, by its nature, filters out the deterministic yearly and seasonal trends, we have generated the kinds of statistical information on hourly solar radiation that can serve as input for analytic design tools\[7,8\].

We have analyzed a climatic database that is broader than most considered in past studies and that spans temperate and tropical climates. We have examined the observation of Liu and Jordan regarding the equivalence of the stationary statistics of solar radiation for individual hours and for daily totals and found it to be accurate for the range of climates considered here. The fact that this equivalence implies that the observations of\[13\] relating to the stationary statistics of daily total radiation also pertain to radiation for individual hours was equally confirmed. Furthermore, we suggest that the simple analytic function proposed in\[13\] for the PDF of daily total radiation be used for individual hours with the input parameters being computed from daily total radiation data only.

Analysis of the sequential behavior of hourly solar radiation reveals strong correlations over long persistence times nearly equal to the entire daylight period. Also, to a good approximation, correlation strengths are not found to depend on time of day, so that estimating a single autocorrelation coefficient for a given lag time representative of the average over all hours of the day appears to be adequate. In addition, careful separation of the effects of lag times of one hour and two hours, via the partial autocorrelation coefficients, reveals that the isolated effect of one-hour lags is by far the predominant, if not the sole, factor. These observations are similar to those made for other assorted meteorological variables\[14-16\]. Finally, we find that the isolated effect of two-hour (or longer) lag times is negligible at .05 significance level, for the vast majority, although by no means for all, of the climatic conditions considered.

Because our analyses are based on discrete hourly intervals, we cannot detect persistence patterns at time lags other than an integer number of hours. This is important since the strongest persistence times could be less than one hour and would remain undetected in our calculations. This implies the need for recording solar radiation (and possibly other meteorolog-
Global horizontal solar radiation on a time scale such as minutes[21]. For both stationary and sequential properties of hourly solar radiation, our findings indicate that universal correlations do not exist. This pertains, for example, to relations between the variance and mean of hourly \( x_t \), as well as to relations between autocorrelation coefficients and average clearness index. Finally, our results are cast in a form that characterizes a given climatic station such that considerable time and expense can be saved in the optimal sizing of solar energy systems.

Acknowledgment—We thank A. Mani for kindly supplying us with radiation data of Indian locations. The critical comments of Y. Zarmi are greatly appreciated.

NOMENCLATURE

- \( d \) index for lag time of persistence effects over consecutive hours
- \( i \) index for hour of day
- \( j \) index for day of month
- \( k \) number of days in month
- \( n \) hourly clearness index for hour \( i \)
- \( K \) daily clearness index
- \( P(Y) \) probability density function for the variable \( Y \)
- \( r \) index for month of year
- \( S \) skewness
- \( t \) index for year
- \( T \) number of years of available data for a particular location
- \( x \), \( k_i/k \)
- \( ar{Y} \) monthly mean value of the variable \( Y \)
- \( \bar{Y} \) grand average of the variable \( Y \) (i.e., average of \( \bar{Y} \) over \( T \) years)
- \( z \) normalized random variable with zero mean and standard deviation of unity
- \( \sigma^2 \) variance
- \( \rho(d) \) autocorrelation coefficient at a lag time of \( d \) hours for hour \( i \)
- \( \phi(d, d) \) partial autocorrelation coefficient at a lag time of \( d \) hours for hour \( i \)

REFERENCES

22. Even a time scale of one month can be too long for filtering out deterministic trends when the latitude is sufficiently high and/or the hour of the day is sufficiently early or late. This is a consequence of the rapid variation of air mass with day of the month in these cases. The qualitatively different nature of hourly global PDF's for high-latitude locations [23] may be a consequence of this effect.